ELSEVIER
Journal of Computational and Applied Mathematics 131 (2001) 161–174
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.nl/locate/cam

Algorithms associated with arithmetic, geometric and harmonic means and integrable systems

Yoshimasa Nakamura *

Department of Informatics and Mathematical Science, Graduate School of Engineering Science, Osaka University, Machikaneyama-cho 1-3, Toyonaka, Osaka 560-8531, Japan

Received 13 September 1999; received in revised form 24 January 2000

Abstract

Gauss' algorithm for arithmetic–geometric mean (AGM) can be regarded as a discrete-time integrable dynamical system having an elliptic theta function solution and a conserved quantity. In this paper we consider algorithms associated with arithmetic, geometric and harmonic means from a viewpoint of integrable systems. First, a max-plus limit and its inverse limit of the AGM algorithm are discussed. These mean operations are shown to be connected to each other by the max-plus limit. Secondly, continous-time dynamical systems associated with the arithmetic–harmonic mean (AHM) algorithm are found. Thirdly, it is shown that the AHM algorithm in indefinite case has a chaotic dynamics and is a generator of numbers which obey the Cauchy distribution. Finally, an extension of the AHM algorithm to the space of positive-definite symmetric matrices is considered.

MSC: 58F07; 39A10; 65F30

Keywords: Arithmetic–geometric mean algorithm; Arithmetic–harmonic mean algorithm; Integrable systems; Max-plus limit; Chaotic dynamics; Positive-definite matrices; Information geometry

1. Introduction

In this decade a closed relationship between certain classes of algorithms and integrable systems is widely understood. Here the integrable systems are dynamical systems which can be explicitly solved, in principle, by the quadrature or some other integration techniques (cf. [13]). Such algorithms and integrable systems are classified into the following four cases.

(i) Algorithms whose one step is just time-one evolution of some integrable system, for example, the QR algorithm [24].

* Fax: +81-6-6850-6496.
E-mail address: naka@sigmath.es.osaka-u.ac.jp (Y. Nakamura).

PII: S0377-0427(00)00316-2

(ii) Algorithms which can be regarded as a time discretization in a linear level of some integrable gradient system, for example, the power method [20].
(iii) Algorithms which can be regarded as a time discretization in a tau-function level of some soliton equation, for example, the qd algorithm [19].
(iv) Integrable gradient systems of Lax type which solving matrix eigenvalue problems, for example, the Brockett equation [7].

In this paper a new class of algorithms having an intimate relationship to integrable systems is considered. The recurrence formulae of algorithms can be regarded as discrete-time integrable systems having explicit solutions and conserved quantities. Each algorithm is defined by a pair of mean operations and generates sequences converging to a common limit in quadratic order. The keyword to connect such algorithms to integrable systems is the existence of additional formulae. In the subsequent part of this section Gauss' *arithmetic–geometric mean algorithm* (cf. [6,8]) is reviewed as a typical example.

Let a_0 and b_0 be given positive numbers such that $a_0 > b_0 > 0$. Let us consider the iteration of the arithmetic mean and the geometric mean operations

$$a_{n+1} = \frac{a_n + b_n}{2}, \quad b_{n+1} = \sqrt{a_n b_n}, \quad n = 0, 1, 2, \ldots, \tag{1}$$

which generates sequences $\{a_n\}_{n=0,1,2,\ldots}$ and $\{b_n\}_{n=0,1,2,\ldots}$. Then it can be easily checked that $b_n < b_{n+1} < \cdots < a_{n+1} < a_n$. The sequences converge to a common limit, say, $M \equiv M(a_0, b_0)$ named the *arithmetic–geometric mean* (AGM) of a_0 and b_0. The complete elliptic integral of the first kind

$$I(a_0, b_0) \equiv \int_0^{\pi/2} \frac{\mathrm{d}\phi}{\sqrt{a_0^2 \cos^2 \phi + b_0^2 \sin^2 \phi}}$$

appears, for example, in computation of period of the simple pendulum. There is an important formula $I(a_0, b_0) = \pi/(2M)$, where M is the AGM. The proof is given by successively applying the Landen transformation to the integral to show $I(a_0, b_0) = I(a_1, b_1) = \cdots = I(M, M)$. The AGM algorithm is known as an excellent algorithm for computing the complete elliptic integral of the first kind (cf. [1, p. 720]).

For the elliptic theta-function

$$\Theta(z;\tau) \equiv \sum_{k \in \mathbf{Z}} \exp(2\pi \mathrm{i} kz + \pi \mathrm{i} k^2 \tau), \quad z, \tau \in \mathbf{C}, \ \operatorname{Im} \tau > 0, \ \mathrm{i} \equiv \sqrt{-1},$$

we set $\Theta_3(\tau) \equiv \Theta(0;\tau)$ and $\Theta_4(\tau) \equiv \Theta(\frac{1}{2};\tau)$. Then we can prove an additional formula

$$\Theta_3(2\tau)^2 = \frac{\Theta_3(\tau)^2 + \Theta_4(\tau)^2}{2}, \quad \Theta_4(2\tau)^2 = \Theta_3(\tau)\Theta_4(\tau) \tag{2}$$

(cf. [25]). Let τ be pure imaginary, and then $q \equiv \exp(\mathrm{i}\pi\tau)$ becomes a real number satisfying $0 < q < 1$ and both $\Theta_3(\tau)$ and $\Theta_4(\tau)$ take a real value. Hence, the nth terms of the AGM algorithm

(1) are expressed by

$$a_n = M\Theta_3(2^n\tau)^2, \quad b_n = M\Theta_4(2^n\tau)^2, \quad n = 0, 1, 2, \ldots \ . \tag{3}$$

Here the pure imaginary number τ is uniquely determined by a_0 and b_0.

Viewing the above fact from integrable systems, the AGM algorithm is a discrete-time integrable system having the elliptic theta-function solution and the conserved quantity $I(a_n, b_n)$. Indeed, Deift–Tomei–Li–Previato pointed out that, by changing the discrete variable n with a continuous variable t, $(a(t),\ b(t)) = (M\Theta_3(2^t\tau)^2,\ M\Theta_4(2^t\tau)^2)$ is an integrable Hamiltonian flow over $\{(a, b) \in \mathbb{R}^2 | a > b > 0\}$, where $H(a, b) \equiv I(a, b)$ is a Hamiltonian (cf. [9]). Theta function (or, quasi-periodic) solutions of soliton equations have been known so far (cf. [26]). Additional formulae with respect to the parameter z appear there. It is to be noted that a *time evolution* with respect to the other parameter τ emerges in the AGM algorithm.

Substituting $c_n \equiv 1/a_n$ and $d_n \equiv 1/b_n$ into the AGM algorithm (1) we obtain

$$c_{n+1} = \frac{2c_n d_n}{c_n + d_n}, \quad d_{n+1} = \sqrt{c_n d_n}. \tag{4}$$

The first of (4) is the harmonic mean of the positive numbers c_n and d_n. Hence the sequences generated by the *harmonic–geometric mean algorithm* (4) starting from c_0, d_0 converge to the common limit $2I(1/c_0, 1/d_0)/\pi$ quadratically. Moreover, a matrix AGM algorithm was introduced in [23] for computing the exponential and logarithm of a given matrix.

In the following sections we discuss the AGM algorithm (1) and the arithmetic–harmonic mean (AHM) algorithm from a viewpoint of integrable systems.

In Section 2, we introduce the max-plus limit of the AGM algorithm. Under this limit the addition changes to the max operation and the multiplication goes to the addition, respectively. Hence, the AGM algorithm turns to the max-arithmetic mean algorithm. On the other hand, by taking the inverse max-plus limit of the geometric mean, we present a hierarchy of the log–geometric means.

In Section 3, it is proved that sequences defined by the AHM algorithm converge in a quadratic order to the geometric mean $N \equiv \sqrt{\alpha_0\beta_0}$ of the positive initial value $\alpha_0 > 0$, $\beta_0 > 0$. The AHM algorithm of positive case is also regarded as a discrete-time integrable system admitting a hyperbolic function solution and a conserved quantity. Two types of continuous-time integrable systems associated with the AHM algorithm are found. One is an integrable Hamiltonian system of Deift–Tomei–Li–Previato type. The other is an integrable system derived from a matrix additional formula of coth by taking a continuous limit.

The AHM algorithm for the indefinite initial value $\alpha_0 > 0$, $\beta_0 < 0$ is discussed in Section 4. The recurrence formula has a trigonometric function solution and a conserved quantity. However, the dynamical system behaves chaotic. We prove that the AHM algorithm of indefinite case is conjugate to a chaotic dynamical system named the Bernoulli shift and is a solvable chaotic system. The corresponding invariant measure and a positive Lyapunov exponent are found explicitly.

Finally in Section 5, we consider three types of geodesics on the space of $m \times m$ real positive-definite symmetric matrices $\mathrm{PD}(m)$ and their relationship to arithmetic, geometric and harmonic means of positive-definite matrices. Here $\mathrm{PD}(m)$ is known as a Riemannian manifold being endowed with an information geometric structure. A matrix extension of the AHM algorithm to $\mathrm{PD}(m)$ is presented. Sequences of positive-definite matrices are shown to converge to a common limit quadratically.

2. Max-plus limit of the AGM algorithm

The *ultra-discrete limit*, so called, is a discretization process starting from a continuous-time dynamical system to a time evolution on a discrete set S through a suitable time discretization of the dynamical system (cf. [27]). The discrete-time system and the time evolution on S is related by the following mapping. For positive numbers a and b we associate $\exp(A/\varepsilon)$ and $\exp(B/\varepsilon)$, respectively, where $\varepsilon > 0$. Then the mapping $\Rightarrow$ is defined by taking a "ε loglimit"

$$\begin{aligned} a+b &\Rightarrow \lim_{\varepsilon\to 0+} \varepsilon \log\left(\exp\frac{A}{\varepsilon} + \exp\frac{B}{\varepsilon}\right) = \max(A,B), \\ ab &\Rightarrow \lim_{\varepsilon\to 0+} \varepsilon \log\left(\exp\frac{A}{\varepsilon} \exp\frac{B}{\varepsilon}\right) = A+B, \end{aligned} \tag{5}$$

namely, the addition $a+b$ changes to $\max(A,B)$ and the multiplication ab goes to the addition $A+B$.

Recently, various soliton cellar-automata are found from soliton equations in terms of the ultra-discrete limit (cf. [28]). *Inverse ultra-discrete limit* is a process for finding underlying continuous-time dynamical systems from cellar automata. Hence the ultra-discrete limit and its inverse cannot be taken twice. While we can take the mapping (5) and its inverse many times only for discrete-time dynamical systems. This is the reason why we name (5) the *max-plus limit*. Remark that the mapping (5) is fundamental in the theory of discrete event systems [5].

Now we consider the max-plus limit of the AGM algorithm (1). Replacing a_n and b_n in (1) with $\exp(A_n/\varepsilon)$ and $\exp(B_n/\varepsilon)$, respectively, we obtain through the max-plus limit

$$A_{n+1} = \max(A_n, B_n), \quad B_{n+1} = \frac{A_n + B_n}{2}, \quad n = 0,1,2,\dots . \tag{6}$$

This recurrence formula generates new sequences $\{A_n\}_{n=0,1,2,\dots}$ and $\{B_n\}_{n=0,1,2,\dots}$. It follows from $a_n > b_n > 0$ that $A_n > B_n$. The nth term is then expressed by

$$A_n = A_0, \quad B_n = \left(1 - \frac{1}{2^n}\right) A_0 + \frac{1}{2^n} B_0, \quad n = 0,1,2,\dots .$$

Since $A_{n+1} - B_{n+1} = (A_n - B_n)/2$, the sequence $\{B_n\}_{n=0,1,2,\dots}$ converges to the limit $A_0 = \max(A_0, B_0)$ in linear order. We note that the max-plus limit of the sequence of geometric means $\{b_n\}_{n=0,1,2,\dots}$ is the geometric progression $\{B_n\}_{n=0,1,2,\dots}$. The new algorithm (6) should be called the *max-arithmetic mean algorithm*. The max operation $\max(A_n, B_n)$ gives a rather rough mean of A_n and B_n. Combining (1) with (6) we see

Proposition 1. *The geometric mean, the arithmetic mean and the max mean operations are related by the max-plus limit* (5) *as follows,*

geometric mean $\Rightarrow$ arithmetic mean $\Rightarrow$ max mean.

Next, we calculate the *inverse max-plus limit* of the AGM algorithm (1). Setting $a_n \equiv \varepsilon \log a'_n$ and $b_n \equiv \varepsilon \log b'_n$ with $a'_n > b'_n > 1$ and taking a limit $\varepsilon \to 0+$, we derive

$$a'_{n+1} = \sqrt{a'_n b'_n}, \quad b'_{n+1} = \exp\sqrt{\log a'_n \log b'_n}, \quad n = 0,1,2,\dots \tag{7}$$

from (1). Conversely, we can recover (1) from (7) by using (5). The sequences $\{a'_n\}_{n=0,1,2,\dots}$ and $\{b'_n\}_{n=0,1,2,\dots}$ converge quadratically. Let us call the second of (7) as the *log–geometric mean* of the positive numbers a'_n and b'_n.

To take the inverse max-plus limit of (7) again we introduce a''_n and b''_n by $a'_n \equiv \varepsilon \log a''_n$ and $b'_n \equiv \varepsilon \log b''_n$ providing that $a''_n > b''_n > e$. Then the recurrence formulae

$$\begin{aligned} a''_{n+1} &= \exp\sqrt{\log a''_n \log b''_n}, \\ b''_{n+1} &= \exp(\exp\sqrt{\log(\log a''_n)\log(\log b''_n)}), \quad n = 0,1,2,\dots \end{aligned} \tag{8}$$

are obtained. We name the second operation of (8) the *double log–geometric mean* of a''_n and b''_n. It follows from (1), (7), (8) that

Proposition 2. *The geometric mean, the log–geometric mean and the double log–geometric mean operations are related by the max-plus limit* (5) *as follows*:

$$\cdots \Rightarrow \textit{double log–geometric mean} \Rightarrow \textit{log–geometric mean} \Rightarrow \textit{geometric mean} \Rightarrow \cdots.$$

There is an inequality among these means. Let $a > b > e$. Then

$$\begin{aligned} a = \max(a,b) &> \frac{a+b}{2} \\ &> \sqrt{ab} \\ &> \exp\sqrt{\log a \log b} \\ &> \exp(\exp\sqrt{\log(\log a)\log(\log b)}) \\ &\ \vdots \\ &> \min(a,b) = b. \end{aligned} \tag{9}$$

Here some new mean operations and algorithms are presented by the max-plus limit and its inverse starting from the AGM algorithm (1).

3. The AHM algorithm

In this section we discuss an iteration of the arithmetic mean and the harmonic mean operations of given two positive numbers α_0 and β_0 such that $\alpha_0 > \beta_0 > 0$. The recurrence formulae

$$\alpha_{n+1} = \frac{\alpha_n + \beta_n}{2}, \quad \beta_{n+1} = \frac{2\alpha_n\beta_n}{\alpha_n + \beta_n}, \quad \alpha_0 > \beta_0 > 0,\ n = 0,1,2,\dots \tag{10}$$

generate the sequences $\{\alpha_n\}_{n=0,1,2,\dots}$ and $\{\beta_n\}_{n=0,1,2,\dots}$. It is easy to show $\beta_n < \beta_{n+1} < \alpha_{n+1} < \alpha_n$. Let $N = N(\alpha_0, \beta_0)$ be the common limit of $\{\alpha_n\}_{n=0,1,2,\dots}$ and $\{\beta_n\}_{n=0,1,2,\dots}$. Let us call (10) the *AHM algorithm*. Noting that

$$\alpha_{n+1} - \beta_{n+1} = \frac{(\alpha_n - \beta_n)^2}{4\alpha_{n+1}} < \frac{1}{4\beta_0}(\alpha_n - \beta_n)^2,$$

we see that the convergence rate is quadratic. From definition (10) $\alpha_{n+1}\beta_{n+1} = \alpha_n\beta_n$ for any n. The common limit is then expressed as $N = \sqrt{\alpha_0\beta_0}$ which is just the geometric mean of the initial value α_0, β_0. The mapping $\alpha = N\coth(\sigma)$, $\beta = N\tanh(\sigma)$ gives a bijection from $\{(\alpha,\beta)\,|\,\alpha > \beta > 0\}$ to $\{(N,\sigma)\,|\,N > 0,\ \sigma > 0\}$.

There is an additional formula for the hyperbolic functions coth and tanh,

$$\coth(2\sigma) = \frac{\coth(\sigma)+\tanh(\sigma)}{2}, \quad \tanh(2\sigma) = \frac{2\coth(\sigma)\tanh(\sigma)}{\coth(\sigma)+\tanh(\sigma)}. \tag{11}$$

The n-term of the AHM algorithm is then

$$\alpha_n = N\coth(2^n\sigma), \quad \beta_n = N\tanh(2^n\sigma), \quad n = 0,1,2,\ldots, \tag{12}$$

where the parameters N and σ are determined uniquely from the given initial values α_0, β_0 as follows:

$$N = \sqrt{\alpha_0\beta_0}, \quad \sigma = \tanh^{-1}\sqrt{\frac{\beta_0}{\alpha_0}}.$$

We conclude that

Theorem 3. *The AHM algorithm with $\alpha_0 > \beta_0 > 0$ is a discrete-time integrable system having the hyperbolic function solution* (12) *and the conserved quantity $\sqrt{\alpha_n\beta_n}$.*

Next we proceed to find continuous-time integrable systems behind the AHM algorithm. One answer is an integrable Hamiltonian system derived in the same manner as that of Deift–Tomei–Li–Previato (cf. [9]). Let us change the discrete variable n in (12) into the continuous variable t as follows:

$$\alpha(t) \equiv N\coth(2^t\sigma), \quad \beta(t) \equiv N\tanh(2^t\sigma), \quad t \geqslant 0. \tag{13}$$

The functions $\alpha(t)$ and $\beta(t)$ satisfy the differential equations

$$\frac{\mathrm{d}\alpha(t)}{\mathrm{d}t} = \log 2 \times 2^t\sigma\sqrt{\frac{\alpha}{\beta}}(\beta-\alpha), \quad \frac{\mathrm{d}\beta(t)}{\mathrm{d}t} = \log 2 \times 2^t\sigma\sqrt{\frac{\beta}{\alpha}}(\alpha-\beta).$$

We consider the Poisson bracket defined by the symplectic 2-form

$$\omega = \frac{1}{\log 2 \times 2^{t+1}\sigma(\alpha-\beta)}\,\mathrm{d}\alpha\wedge\mathrm{d}\beta.$$

Then the differential equations take the following nonautonomous Hamiltonian form:

$$\frac{\mathrm{d}\alpha(t)}{\mathrm{d}t} = \log 2 \times 2^{t+1}\sigma(\beta-\alpha)\frac{\partial H}{\partial\beta}, \quad \frac{\mathrm{d}\beta(t)}{\mathrm{d}t} = -\log 2 \times 2^{t+1}\sigma(\beta-\alpha)\frac{\partial H}{\partial\alpha} \tag{14}$$

with respect to the Hamiltonian $H(\alpha,\beta) \equiv \sqrt{\alpha\beta}$, where $\beta(t)-\alpha(t)$ is not zero for any $t \geqslant 0$. Conversely, the sequences $\{\alpha_n\}_{n=0,1,2,\ldots}$ and $\{\beta_n\}_{n=0,1,2,\ldots}$ lie on the continuous flow of (14).

Proposition 4 (Deift–Tomei–Li–Previato)**.** *$(\alpha(t),\beta(t))$ is an integrable Hamiltonian flow equipped with the symplectic structure. The Hamiltonian is just $H(\alpha,\beta) = \sqrt{\alpha\beta}$.*

The other answer is a continuous limit of the AHM algorithm (10). We rewrite the additional formula (11) of coth in a matrix form found in [4],

$$\begin{aligned}\coth(2\sigma) &= \frac{\coth^2(\sigma)+1}{2\coth(\sigma)}\\ &= -\frac{\begin{vmatrix}\coth(\sigma) & \coth(\sigma-\mu)\\ \coth(\sigma+\mu) & \coth(\sigma)\end{vmatrix}}{\begin{vmatrix}\coth(\sigma) & \coth(\sigma-\mu) & 1\\ \coth(\sigma+\mu) & \coth(\sigma) & 1\\ 1 & 1 & 0\end{vmatrix}}\\ &= \frac{\begin{vmatrix}\coth(\sigma-\mu) & \coth(\sigma)-\coth(\sigma-\mu)\\ \coth(\sigma)-\coth(\sigma-\mu) & \coth(\sigma+\mu)-2\coth(\sigma)+\coth(\sigma-\mu)\end{vmatrix}}{\coth(\sigma+\mu)-2\coth(\sigma)+\coth(\sigma-\mu)},\end{aligned} \tag{15}$$

where μ is an arbitrary real number. Multiplying to both the numerator and the denominator by $1/\mu^2$ and taking a limit $\mu \to 0$, we derive a new additional formula

$$\coth(2\sigma) = \frac{\begin{vmatrix}\coth(\sigma) & \dfrac{\mathrm{d}\coth(\sigma)}{\mathrm{d}\sigma}\\ \dfrac{\mathrm{d}\coth(\sigma)}{\mathrm{d}\sigma} & \dfrac{\mathrm{d}^2\coth(\sigma)}{\mathrm{d}\sigma^2}\end{vmatrix}}{\mathrm{d}^2\coth(\sigma)/\mathrm{d}\sigma^2} \tag{16}$$

which includes derivatives. It is shown that

Theorem 5. *A continuous limit of the AHM algorithm* (10) *is the delay-differential equation*

$$u(2t)\frac{\mathrm{d}^2u(t)}{\mathrm{d}t^2} = u(t)\frac{\mathrm{d}^2u(t)}{\mathrm{d}t^2} - \left(\frac{\mathrm{d}u(t)}{\mathrm{d}t}\right)^2 \tag{17}$$

which has explicit solutions $\coth(t)$ *and* $\cot(t)$.

The dynamical system (17) is rather different from (14). Note the positivity of $u(t) = \coth(t)$ for $t \geqslant 0$. The AHM algorithm (10) is, conversely, a time discretization of (17) which keeps the positivity. Hence we say (10) as an *integrable discretization* of (17). Here the terminology, integrable discretization, is defined recently in [20]. The other solution $\cot(t)$ is associated with an indefinite initial value $\alpha_0 > 0$ and $\beta_0 < 0$ and violates the positivity, which will be discussed in the next section.

Finally, in this section we remark on the matrix additional formula (15). Let us fix μ in (15) as $\mu = \sigma/l$ for some natural number l. Set $\alpha_l \equiv \coth(\sigma)$ and $\alpha_{l\pm1} \equiv \coth(\sigma \pm \mu)$, and consequently, $\alpha_{2l} \equiv \coth(2\sigma)$. Let us introduce the mapping

$$\alpha_n = f(\alpha_{n-1}).$$

Then the additional formula for coth is expressed as

$$\alpha_{2l} = \frac{\begin{vmatrix}\alpha_{l-1} & f(\alpha_{l-1})\\ f(\alpha_{l-1}) & f(f(\alpha_{l-1}))\end{vmatrix}}{f(f(\alpha_{l-1})) - 2f(\alpha_{l-1}) + \alpha_{l-1}}. \tag{18}$$

Eq. (18) is known as the recurrence formula of the Steffensen iteration for finding one real solution of the nonlinear equation $x - f(x) = 0$ (cf. [22, p. 241]). The existence of an additional formula representation reflects the quadratic convergence of the Steffensen iteration.

4. The AHM algorithm: indefinite case

In this section, we restrict ourselves to the AHM algorithm

$$\alpha_{n+1} = \frac{\alpha_n + \beta_n}{2}, \quad \beta_{n+1} = \frac{2\alpha_n\beta_n}{\alpha_n + \beta_n}, \quad \alpha_0 > 0,\ \beta_0 < 0,\ n = 0, 1, 2, \ldots \tag{19}$$

of an indefinite initial value case. An additional formula

$$\cot(2\sigma) = \frac{\cot(\sigma) + (-\tan(\sigma))}{2}, \quad -\tan(2\sigma) = \frac{2\cot(\sigma)(-\tan(\sigma))}{\cot(\sigma) + (-\tan(\sigma))} \tag{20}$$

of the trigonometric functions plays an important role here. There is a bijection from the set $\{(\alpha, \beta) \mid \alpha > 0, \beta < 0\}$ to $\{(N, \sigma) \mid N > 0, 0 < \sigma < \pi/2\}$ defined by

$$N = \sqrt{-\alpha_0\beta_0}, \quad \sigma = \tan^{-1}\sqrt{-\frac{\beta_0}{\alpha_0}}.$$

Here N is the geometric mean of α_0 and $-\beta_0$. The nth term of the recurrence formula is given by

$$\alpha_n = N\cot(2^n\sigma), \quad \beta_n = -N\tan(2^n\sigma), \quad n = 0, 1, 2, \ldots\ . \tag{21}$$

The AHM algorithm (19) is a discrete-time solvable dynamical system which has an explicit solution (21) and the conserved quantity $N = \sqrt{-\alpha_n\beta_n}$. However, α_n and β_n do not converge as $n \to \infty$ in general. This is a sharp contrast with the positive case where $\alpha_0 > \beta_0 > 0$. Fig. 1 indicates the behaviour of solution α_n with $\alpha_0 = 0.5$, $\beta_0 = -2$.

Next, we consider a chaotic behaviour which appears in Fig. 1. Eliminating β_n by using the conserved quantity $N = \sqrt{-\alpha_n\beta_n}$ we derive from (19)

$$\alpha_{n+1} = g(\alpha_n) \equiv \frac{1}{2}\left(\alpha_n - \frac{N^2}{\alpha_n}\right). \tag{22}$$

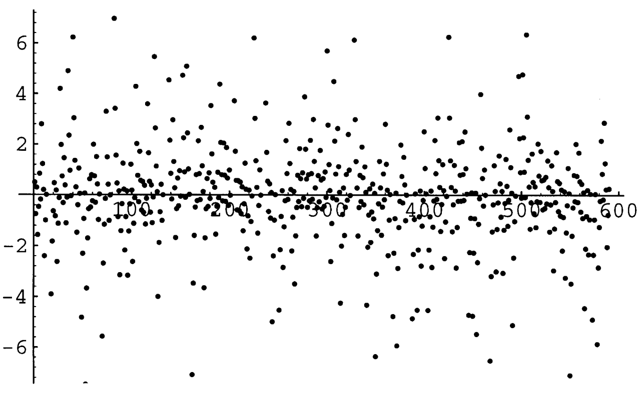

Fig. 1. The AHM algorithm with $\alpha_0 = 0.5$, $\beta_0 = -2$.

Let us rewrite (22) by the change of variables

$$\alpha_n = -N\cot(\pi\omega_n) \quad (0 < \omega_n < 1)$$

to $\cot(\pi\omega_{n+1}) = \cot(2\pi\omega_n)$. It follows that ω_n satisfies

$$\omega_{n+1} = \begin{cases} 2\omega_n & 0 < \omega_n \leqslant \frac{1}{2}, \\ 2\omega_n - 1 & \frac{1}{2} < \omega_n < 1. \end{cases} \tag{23}$$

Eq. (23) is the dynamical system of the Bernoulli shift (cf. [11]). The solvability enables us to find an invariant measure and the exact Lyapunov exponent for (19). It is to be noted that the resulting invariant measure

$$\rho(x) = \frac{N}{\pi(N^2 + x^2)} \quad (-\infty < x < \infty)$$

is congruent with the density function of a Cauchy distribution. This fact explains the distribution of points $\{\alpha_n\}_{n=0,1,2,\ldots}$ in Fig. 1. The AHM algorithm of indefinite case is shown to be a generator of the random variables which obey the Cauchy distribution having the parameter N. The Lyapunov exponent λ is given by $\lambda = \log 2 > 0$. We have

Theorem 6. *The AHM algorithm of indefinite case* (19) *is a* solvable chaotic system *which is conjugate to the Bernoulli shift* (23).

Recently, Umeno [29] found a class of solvable chaotic systems associated with Lévi's stable distributions including the Cauchy distribution.

5. The AHM algorithm on the space of positive-definite matrices

First in this section we give a differential–geometric interpretation of the AHM algorithm. Let $\mathrm{PD}(m)$ be a set of $m \times m$ real positive-definite symmetric matrices. Such matrices appear in the stability theorem for linear dynamical systems, covariance matrices of stochastic systems, convex optimization problems and so on. Let $\theta = (\theta^i)$ be a coordinates system of $\mathrm{PD}(m)$ such that $\mathrm{PD}(m) \ni P = \theta^i E_i$, where E_i are matrix elements. Then $\mathrm{PD}(m)$ is known as a Riemannian manifold having the Riemannian metric $g_{ij}(\theta) = \mathrm{tr}(P^{-1}E_iP^{-1}E_j)$ (see [21]). There are three type of geodesics due to an information (or, dualistic) geometry structure of $\mathrm{PD}(m)$.

The first is ∇-geodesics. Let P be a point of $\mathrm{PD}(m)$ and X be an element of the tangent space $T_P\mathrm{PD}(m)$. The ∇-connection is determined by the parallel translation $\Pi_C : X \equiv X_{P(t_1)} \mapsto X_{P(t_2)} = X$. The corresponding geodesics equation is simply $\mathrm{d}^2P(t)/\mathrm{d}t^2 = 0$. The geodesics from the point $P(0) = P_1$ to the target $P(1) = P_2$ is $P(t) = P_1 + (P_2 - P_1)t$. We are interested in its midpoint

$$\mathrm{A} \equiv P(\tfrac{1}{2}) = \tfrac{1}{2}(P_1 + P_2). \tag{24}$$

The point A is just the arithmetic mean of matrices P_1 and P_2.

The second is ∇^*-geodesics. The ∇^*-connection corresponds to the parallel translation $\Pi_C^* : X \equiv X_{P(t_1)} \mapsto X_{P(t_2)} = P(t_2)P(t_1)^{-1}XP(t_1)^{-1}P(t_2)$. The dualistic coordinate of P is given by its inverse P^{-1}. Then the geodesics equation $\mathrm{d}^2P^{-1}(t)/\mathrm{d}t^2 = 0$ leads to the geodesics $P(t) = (P_1^{-1} + (P_2^{-1} - P_1^{-1})t)^{-1}$

from P_1 to P_2. The midpoint H is

$$\mathrm{H} \equiv P(\tfrac{1}{2}) = 2(P_1^{-1} + P_2^{-1})^{-1}, \tag{25}$$

which expresses the harmonic mean of P_1 and P_2.

The usual geodesics which comes from the Riemannian connection is

$$P(t) = P_1^{1/2}(\exp(\log(P_1^{-1/2}P_2P_1^{-1/2})t))P_1^{1/2}.$$

The midpoint G of the Riemannian geodesics

$$\mathrm{G} \equiv P(\tfrac{1}{2}) = P_1^{1/2}(P_1^{-1/2}P_2P_1^{-1/2})^{1/2}P_1^{1/2} \tag{26}$$

is related to a geometric mean of the positive-definite matrices P_1 and P_2.

We now design an extension of the AHM algorithm for the positive-definite symmetric matrices Q_0 and R_0 as follows:

$$\begin{aligned} Q_{n+1} &= \tfrac{1}{2}(Q_n + R_n), \\ R_{n+1} &= 2(Q_n^{-1} + R_n^{-1})^{-1}, \quad n = 0, 1, 2, \ldots . \end{aligned} \tag{27}$$

Then the sequences of matrices $\{Q_n\}_{n=0,1,2,\ldots}$ and $\{R_n\}_{n=0,1,2,\ldots}$ are generated by (27). We give two lemmas here.

Lemma 7. *If $Q_0 = I$, then $Q_nR_n = R_nQ_n$, $n = 0, 1, 2, \ldots$.*

Proof. Obviously, $Q_0R_0 = R_0Q_0$. Let us assume that $Q_kR_k = R_kQ_k$ for some integer k. Noting $Q_k^{-1}R_k = R_kQ_k^{-1}$ and $R_k^{-1}Q_k = Q_kR_k^{-1}$ we derive $(Q_k^{-1} + R_k^{-1})(Q_k + R_k) = (Q_k + R_k)(Q_k^{-1} + R_k^{-1})$. It follows that $Q_{k+1}R_{k+1} = R_{k+1}Q_{k+1}$. Lemma 7 is proved by induction. □

Lemma 8. *If $Q_0 = I$, then $Q_{n+1}R_{n+1}^2Q_{n+1} = Q_nR_n^2Q_n$, $n = 0, 1, 2, \ldots$. Thus,*

$$N \equiv Q_nR_n^2Q_n \tag{28}$$

is a conserved quantity of the recurrence formula (27).

Proof. We have

$$\begin{aligned} (Q_n^{-1} + R_n^{-1})^2 &= R_nQ_n^{-1}R_n^{-2}Q_n^{-1}R_n + R_nQ_n^{-1}R_n^{-2} + R_n^{-2}Q_n^{-1}R_n + R_n^{-2} \\ &= (Q_n + R_n)Q_n^{-1}R_n^{-2}Q_n^{-1}(Q_n + R_n) \end{aligned}$$

with the help of Lemma 7. Taking the inverse we prove

$$\begin{aligned} Q_nR_n^2Q_n &= (Q_n + R_n)(Q_n^{-1} + R_n^{-1})^{-2}(Q_n + R_n) \\ &= Q_{n+1}R_{n+1}^2Q_{n+1}. \qquad \square \end{aligned}$$

We have the main theorem.

Theorem 9. *The sequences $\{Q_n\}_{n=0,1,2,\ldots}$ and $\{R_n\}_{n=0,1,2,\ldots}$ with $Q_0 = I$ tend to the common limit $\mathrm{G} = R_0^{1/2}$ in a quadratic order.*

Proof. We first remark the convexity of PD(m), namely, both $Q+R$ and $(Q^{-1}+R^{-1})^{-1}$ are positive-definite for positive-definite matrices Q and R. Socondly, it can be checked that

$$\begin{aligned} Q_{n+1}-R_{n+1} &= \tfrac{1}{2}Q_n(Q_n^{-1}+R_n^{-1}-4(Q_n+R_n)^{-1})R_n \\ &= \tfrac{1}{2}Q_n(R_n^{-1}(Q_n-R_n)-Q_n^{-1}(Q_n-R_n))(Q_n+R_n)^{-1}R_n \\ &= \tfrac{1}{2}Q_nR_n^{-1}(Q_n-R_n)Q_n^{-1}(Q_n-R_n)(Q_n+R_n)^{-1}R_n \\ &= \tfrac{1}{2}(Q_n-R_n)^2(Q_n+R_n)^{-1}, \end{aligned}$$

$$Q_{n+1}-Q_n = \tfrac{1}{2}(R_n-Q_n),$$

$$R_{n+1}-R_n = (Q_n-R_n)(Q_n+R_n)^{-1}R_n, \quad n=0,1,2,\ldots,$$

by using Lemma 7. Hence, the sequences $\{Q_n\}_{n=0,1,2,\ldots}$ and $\{R_n\}_{n=0,1,2,\ldots}$ quadratically converge to a common limit, say, S.

The existence of conserved quantity (Lemma 8) helps us to compute S explicitly. It follows from $Q_0=I$ and (28) that $N=R_0^2$. While, we see that $N=S^4$ by taking the limit $n\to\infty$ in (28). Hence, the sequences $\{Q_n\}_{n=0,1,2,\ldots}$ and $\{R_n\}_{n=0,1,2,\ldots}$ tend to $S=R_0^{1/2}$, the positive-definite square root R_0. □

It is shown that the *matrix AHM algorithm* (27) with $Q_0=I$ provides a method for computing the square root $R_0^{1/2}$ of a given positive-definite symmetric matrix R_0. Algorithm (27) with a special choice of the initial data Q_0 is essentially equivalent to the algorithm introduced by Denman and Beavers Jr. [10]. This algorithm is better than Heron's method

$$A_{n+1} = \tfrac{1}{2}(A_n+R_0A_n^{-1}), \quad A_0=I, \quad n=0,1,2,\ldots \tag{29}$$

for computing $R_0^{1/2}$ in the sense in which the condition numbers of the matrices Q_n, R_n in (27) improve with every iteration. We give a numerical example. Consider the Hilbert matrix

$$H^{(m)}=(h_{ij}), \quad h_{ij}=\frac{1}{i+j-1}, \quad i,j=1,\ldots,m$$

which is a typical example of ill-conditioned positive-definite symmetric matrix (cf. [16]). Let us set $R_0=H^{(4)}$. Applying the matrix AHM algorithm (27) with $Q_0=I$ we see that Q_n "converges" to

$$\begin{pmatrix} 0.9114603871 & 0.33903123 & 0.1927197149 & 0.1309843472 \\ 0.33903123 & 0.3528552107 & 0.2474522474 & 0.1806979358 \\ 0.1927197149 & 0.2474522474 & 0.2411020518 & 0.2085576596 \\ 0.1309843472 & 0.1806979358 & 0.2085576596 & 0.2226032396 \end{pmatrix} \approx R_0^{1/2}$$

for $n=10,11,\ldots$. Here we use REDUCE 3.6. On the other hand, A_{10} and A_{14}, for example, generated by Heron's method (29) are

$$A_{10} = \begin{pmatrix} 0.9114603876 & 0.339031225 & 0.1927197269 & 0.1309843394 \\ 0.3390312219 & 0.3528553023 & 0.2474520267 & 0.1806980792 \\ 0.192719734 & 0.247452032 & 0.2411025704 & 0.2085573224 \\ 0.1309843347 & 0.1806980775 & 0.2085573183 & 0.2226034615 \end{pmatrix},$$

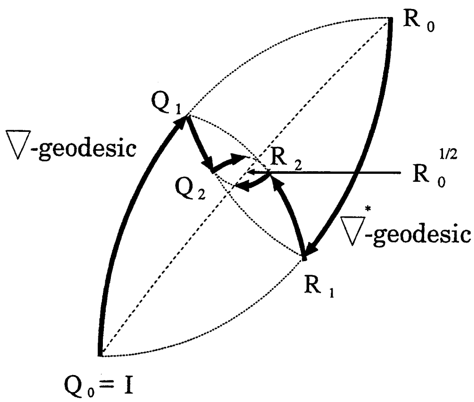

Fig. 2. The matrix AHM algorithm.

$$A_{14} = \begin{pmatrix} 0.9076043572 & 0.3824495113 & 0.08818531537 & 0.1989496125 \\ 0.3368326342 & 0.3776110487 & 0.1878497806 & 0.2194497484 \\ 0.191151168 & 0.2651138346 & 0.1985797924 & 0.2362044118 \\ 0.1297575909 & 0.1945110162 & 0.1753011135 & 0.244225691 \end{pmatrix},$$

respectively. Roundoff error causes unstability.

Finally, in this section we consider the recurrence formula (27) for *arbitrary* positive-definite symmetric matrices Q_0 and R_0. Introduce Q_0' and R_0' by

$$Q_n' \equiv Q_0^{-1/2} Q_n Q_0^{-1/2}, \quad R_n' \equiv Q_0^{-1/2} R_n Q_0^{-1/2}, \quad n = 0, 1, 2, \dots . \tag{30}$$

Obviously, $Q_0' = I$. Theorem 9 implied that $\{Q_n'\}_{n=0,1,2,\dots}$ and $\{R_n'\}_{n=0,1,2,\dots}$ defined by (27) converge to $R_0'^{1/2}$ quadratically. Hence, we see that $\{Q_n\}_{n=0,1,2,\dots}$ and $\{R_n\}_{n=0,1,2,\dots}$ tend to

$$Q_0^{1/2} R_0'^{1/2} Q_0^{1/2} = Q_0^{1/2} (Q_0^{-1/2} R_0 Q_0^{-1/2})^{1/2} Q_0^{1/2}.$$

The right-hand side is just the midpoint of the Riemannian geodesics from Q_0 to R_0. Moreover, the identity

$$Q_0^{1/2} (Q_0^{-1/2} R_0 Q_0^{-1/2})^{1/2} Q_0^{1/2} = R_0^{1/2} (R_0^{-1/2} Q_0 R_0^{-1/2})^{1/2} R_0^{1/2}$$

can be checked. This completes the proof. □

Theorem 10. *The AHM algorithm on the space* PD(m) *of positive-definite symmetric matrices generates sequences* $\{Q_n\}_{n=0,1,2,\dots}$ *and* $\{R_n\}_{n=0,1,2,\dots}$ *which converge quadratically to the midpoint*

$$\mathrm{G} = Q_0^{1/2} (Q_0^{-1/2} R_0 Q_0^{-1/2})^{1/2} Q_0^{1/2} \tag{31}$$

of the Riemannian geodesics from Q_0 *to* R_0.

In other words, it can be said that the matrix AHM algorithm (27) on PD(m) approaches to G along ∇-geodesics and ∇^*-geodesics alternately. Some efficient algorithms having a mathematical structure similar to (27) are known. They are the affine scaling algorithm (cf. [12]) for linear programming problems and the EM algorithm (cf. [3]) for the maximal likelihood estimation. See also a book [2] for information geometry and Refs. [17,18] for integrable systems on information spaces. In Fig. 2, the convergence to $\mathrm{G} = R_0^{1/2}$ of $\{Q_n\}_{n=0,1,2,\dots}$ and $\{R_n\}_{n=0,1,2,\dots}$ defined by the matrix arithmetic–harmonic mean algorithm with $Q_0 = I$ is illustrated.

In operator theory such a common limit as (31) is known as the AHM of positive operators [14,15]. We here give a linear algebraic proof of quadratic convergence of algorithm (27) as well as the existence of the common limit (31).

6. Conclusions

Gauss' AGM algorithm (1) is an additional formula for the elliptic theta-functions. In Section 3 we note that the AHM algorithm (10) is analogously an additional formula for the hyperbolic functions. Both the two algorithms can be regarded as discrete-time integrable systems having conserved quantities. The existence of additional formulae conversely guarantees a quadratic convergence of these algorithms. The positivity of each term of the AGM algorithm enable us to take the max-plus limit (5) of the algorithm in Section 2. A matrix extension of the AHM algorithm (27) on the manifold of positive-definite symmetric matrices is presented in Section 5. Sequences of positive-definite matrices converge to the common limit (31) quadratically. On the other hand, the AHM algorithm of indefinite case (19) is shown to be a chaotic system in Section 4. The algorithm remarkably generates numbers which obey the Cauchy distribution. The positivity seems most crucial in algorithms associated with arithmetic, geometric and harmonic means.

Acknowledgements

The author would like to thank P. Deift, A. Ohara, K. Umeno for stimulating discussions and giving him helpful suggestions. This work is supported in part by Grant-in-Aid for Scientific Research nos. 09440077, 09559011, 10874019 from the Japan Ministry of Education, Sciences, and Culture.

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, Dover, New York, 1972.
[2] S. Amari, Differential–Geometrical Methods in Statistics, Lecture Notes in Statistics, Vol. 28, Springer, Berlin, 1985.
[3] S. Amari, Information geometry of the em and EM algorithms for neural networks, Neural Networks 8 (1995) 1379–1408.
[4] M. Arai, K. Okamoto, Y. Kametaka, Aitken–Steffensen acceleration and a new addition formula for Fibonacci numbers, Proc. Japan Acad. Ser. A 62 (1986) 5–7.
[5] F.L. Baccelli, G. Cohen, G.J. Olsder, J.-P. Quadrat, Synchronization and Linearity. An Algebra for Discrete Event Systems, Wiley, Chichester, 1992.
[6] J.-B. Bost, J.-F. Mestre, Moyenne arithmético-géométrique et périodes des courbes de genre 1 et 2, Gaz. Math. 38 (1988) 36–64.
[7] R.W. Brockett, Dynamical systems that sort lists, diagonalize matrices and solve linear programming problems, Linear Algebra Appl. 146 (1991) 79–91.
[8] D.A. Cox, Gauss and the arithmetic-geometric mean, Notices Amer. Math. Soc. 32 (1985) 147–151.

[9] P. Deift, Continuous versions of some discrete maps or what goes on when the lights go out, J. Anal. Math. 58 (1992) 121–133.
[10] E.D. Denman, A.N. Beavers Jr, The matrix sign function and computations in systems, Appl. Math. Comput. 2 (1976) 63–94.
[11] R.L. Devaney, An Introduction to Chaotic Systems, Benjamin/Cummings, New York, 1986.
[12] I.I. Dikin, Iterative solution of problems of linear and quadratic programming, Sov. Math. Dokl. 8 (1967) 674–675.
[13] B.A. Dubrovin, I.M. Krichber, S.P. Novikov, Integrable systems I, in: V.I. Arnol'd, S.P. Novikov (Eds.), Dynamical Systems IV, Encyclopedia Math. Sci., Vol. 4, Springer, Berlin, 1990, pp. 173–280.
[14] J. Fujii, Arithemetico–geometric mean of operators, Math. Japonica 23 (1979) 667–669.
[15] W.L. Green, T.D. Morley, Operator means and matrix functions, Linear Algebra Appl. 137/138 (1990) 453–465.
[16] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985.
[17] Y. Nakamura, Completely integrable gradient systems on the manifolds of Gaussian and multinomial distributions, Japan J. Indust. Appl. Math. 10 (1993) 179–189.
[18] Y. Nakamura, A tau-function for the finite Toda molecule, and information spaces, in: Y. Maeda, H. Omori, A. Weinstein (Eds.), Contemporary Mathematics, Vol. 179, American Mathematical Society, Providence, RI, 1994, pp. 205–211.
[19] Y. Nakamura, Calculating Laplace transforms in terms of the Toda molecule, SIAM J. Sci. Comput. 20 (1999) 306–317.
[20] Y. Nakamura, K. Kajiwara, H. Shiotani, On an integrable discretization of the Rayleigh quotient gradient system and the power method with a shift, J. Comput. Appl. Math. 96 (1998) 77–90.
[21] A. Ohara, N. Suda, S. Amari, Dualistic differential geometry of positive definite matrices and its applications to related problems, Linear Algebra Appl. 247 (1996) 31–53.
[22] A. Ostrowski, Solution of Equations and Systems of Equations, Academic Press, New York, 1966.
[23] E.U. Stickel, Fast computation of matrix exponential and logarithm, Analysis 5 (1985) 163–173.
[24] W.W. Symes, The QR algorithm and scattering for the finite nonperiodic Toda lattice, Physica 4D (1982) 275–280.
[25] J. Tannery, J. Molk, Éléments de la Théorie des Fonctions Elliptiques II, Chelsea, New York, 1972.
[26] M. Toda, Theory of Nonlinear Lattices, Springer, Berlin, 1981.
[27] T. Tokihiro, D. Takahashi, J. Matsukidaira, J. Satsuma, From soliton equations to integrable cellular automata through a limiting procedure, Phys. Rev. Lett. 76 (1996) 3247–3250.
[28] S. Tsujimoto, R. Hirota, Ultradiscrete KdV equation, J. Phys. Soc. Japan 67 (1998) 1809–1810.
[29] K. Umeno, Superposition of chaotic processes with convergence to Lévy's stable law, Phys. Rev. E 58 (1998) 2644–2647.